Yuting Wei

Statistics & Data Science
307 Academic Research Building
The Wharton School, University of Pennsylvania

Email: ytwei@wharton.upenn.edu
Homepage: `https://yutingwei.github.io/`

Employment

University of Pennsylvania
- Assistant Professor, Statistics and Data Science, The Wharton School, 07/2021 - present
- Affiliated Faculty, Applied Mathematics and Computational Science (AMCS), 01/2022 - present

Carnegie Mellon University
- Assistant Professor, Statistics and Data Science, 08/2019 - 06/2021

Stanford University
- Stein Fellow/Lecturer, Department of Statistics, 09/2018 - 07/2019

Education

University of California, Berkeley, Ph.D. in Statistics, 2013 - 2018
- Thesis: A geometric perspective on some topics in statistical learning
- Advisors: Martin Wainwright, Adityanand Guntuboyina
- Thesis committee: Martin Wainwright, Adityanand Guntuboyina, Peter Bickel, Venkat Anantharam

Peking University, Bachelor of Science (B.S.) Applied Mathematics, 2009 - 2013
- Advisor: Jinzhu Jia, Peking University

Peking University, Bachelor of Arts (B.A.) Economics, 2010 - 2013

Short-term Appointments

Simons Institute for the Theory of Computing, University of California, Berkeley
Visiting scientist, 08/2021 - 12/2021
Institute for Mathematical Research (FIM), ETH Zürich
Visiting researcher, 09/2015 - 12/2015
Department of Mathematics, City University of Hong Kong
Visiting researcher, 08/2012 - 11/2012

Research Interests

- High-dimensional and nonparametric statistics
- Statistical machine learning
- Reinforcement learning
- Applications in bioinformatics and statistical genomics

Selected Honors & Awards

Visiting International Professor Fellowship, Ruhr-Universität Bochum, 2023

NSF CAREER Award, 2022

INFORMS George Nicholson Award, Finalist, 2021

Stein Fellowship, Stanford University, 2019

Erich L. Lehmann Citation, University of California, Berkeley, 2018

Oberwolfach Leibniz Graduate Student Grant, 2017

Outstanding Graduate of Peking University, 2013

Outstanding Winner of Annual Peking University Research Competition (top 10/5000), 2013

Best Overall Winner in 2nd Machine Learning Competition in Immunology, 2012

President Research Fund, Peking University, 2011-2012

Preprints & Submitted Manuscripts

P1. Gen Li, Wei Fan, **Yuting Wei**, "Approximate message passing from random initialization with applications to Z_2 synchronization," arxiv:2302.03682, 2023.

P2. Gen Li, **Yuting Wei**, "A non-asymptotic framework for approximate message passing in spiked models," arxiv:2208.03313, 2022.

P3. Pratik Patil, Arun Kumar Kuchibhotla, **Yuting Wei**, Alessandro Rinaldo, "Mitigating multiple descents: A model-agnostic framework for risk monotonization," arxiv:2205.12937, 2022.

P4. Gen Li, Laixi Shi, Yuxin Chen, Yuejie Chi, **Yuting Wei**, "Settling the Sample Complexity of Model-Based Offline Reinforcement Learning," arxiv:2204.05275, 2022.

P5. Yue Li, **Yuting Wei**, "Minimum ℓ_1 Interpolators: Precise Asymptotics and Multiple Descent," arxiv:2110.09502, 2021.

P6. Shicong Cen, **Yuting Wei**, Yuejie Chi, "Fast Policy Extragradient Methods for Competitive Games with Entropy Regularization," arXiv:2105.15186, 2021.

P7. Michael Celentano, Andrea Montanari, **Yuting Wei** (alphabetical order), "The Lasso with General Gaussian Designs with Applications to Hypothesis Testing," arXiv:2007.13716, 2020.

P8. Gen Li, Changxiao Cai, Yuxin Chen, **Yuting Wei**, Yuejie Chi, "Is Q-Learning Minimax Optimal? A Tight Sample Complexity Analysis," arXiv:2102.06548, under major revision, *Operations Research*, 2021.

P9. Jingyan Wang, Ivan Stelmakh, **Yuting Wei**, Nihar B. Shah, "Debiasing Evaluations That are Biased by Evaluations," under minor revision, *Journal of Machine Learning Research*, 2021.

P10. Gen Li, **Yuting Wei**, Yuejie Chi, and Yuxin Chen, "Breaking the Sample Size Barrier in Model-Based Reinforcement Learning with a Generative Model," arXiv:2005.12900, under major revision, *Operations Research*, 2020.

Journal Articles

J1. Gen Li, **Yuting Wei**, Yuejie Chi, Yuxin Chen, "Softmax Policy Gradient Methods Can Take Exponential Time to Converge," arXiv:2102.11270, accepted to *Mathematical Programming*, 2022.

J2. Zhimei Ren, **Yuting Wei**, Emmanuel J. Candès, "Derandomizing Knockoffs," arxiv:2012.02717, *Journal of the American Statistical Association (JASA)*, 2021.

J3. Shicong Cen, Chen Cheng, Yuxin Chen, **Yuting Wei** and Yuejie Chi, "Fast Global Convergence of Natural Policy Gradient Methods with Entropy Regularization," *Operations Research*, vol. 70, no. 4, pp. 2563-2578, 2022. **INFORMS George Nicholson Paper Award Finalist**.

J4. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Sample Complexity of Asynchronous Q-Learning: Sharper Analysis and Variance Reduction," *IEEE Transactions on Information Theory*, vol. 68, no. 1, pp. 448-473, 2022.

J5. Minshi Peng, Yue Li, Brie Wamsley, **Yuting Wei**, Kathryn Roeder, "Integration and transfer learning of single-cell transcriptomes via cFIT," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 118, no. 10, 2021.

J6. Minshi Peng, Brie Wamsley, Andrew Elkins, Daniel M Geschwind, **Yuting Wei**, Kathryn Roeder, "Cell Type Hierarchy Reconstruction via Reconciliation of Multi-resolution Cluster Tree," *Nucleic Acids Research*, vol. 49, no. 16, 2021.

J7. Chen Cheng, **Yuting Wei**, and Yuxin Chen, "Tackling Small Eigen-gaps: Fine-Grained Eigenvector Estimation and Inference under Heteroscedastic Noise," *IEEE Transactions on Information Theory*, vol. 67, no. 11, pp. 7380-7419, 2021.

J8. **Yuting Wei** and Martin Wainwright, "The Local Geometry of Testing in Ellipses: Tight Control via Localized Kolmogorov Widths," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5110-5129, 2020.

J9. **Yuting Wei**, Billy Fang, and Martin Wainwright, "From Gauss to Kolmogorov: Localized Measures of Complexity for Ellipses," *Electronic Journal of Statistics,* vol. 14, no. 2, pp. 2988-3031, 2020.

J10. **Yuting Wei**, Martin Wainwright and Adityanand Guntuboyina, "The Geometry of Hypothesis Testing over Convex Cones: Generalized Likelihood Tests and Minimax Radii," *Annals of Statistics,* vol. 47, no. 2, pp. 994-1024, 2019.

J11. **Yuting Wei***, Fanny Yang* and Martin Wainwright, "Early Stopping for Kernel Boosting Algorithms: A General Analysis with Localized Complexities," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6685-6703, 2019.

J12. Tony Cai, Adityanand Guntuboyina and **Yuting Wei** (alphabetical order), "Adaptive Estimation of Planar Convex Sets," *Annals of Statistics,* vol. 46, no. 3, pp. 1018-1049, 2018.

J13. Wen-Jun Shen*, **Yuting Wei***, Xin Guo*, Stephen Smale, Hau-San Wong and Shuaicheng Li, "MHC Binding Prediction with KernelRLSpan and Its Variations," *Journal of Immunological Methods,* vol. 406, pp. 10-20, 2014.

Refereed Conference Proceedings

C1. Gen Li, Yuejie Chi, **Yuting Wei**, Yuxin Chen, "Minimax-optimal multi-agent RL in Markov games with a generative model," *Neural Information Processing Systems (NeurIPS)*, **oral presentation**, 2022.

C2. Laixi Shi, Gen Li, **Yuting Wei**, Yuxin Chen, Yuejie Chi, "Pessimistic Q-Learning for offline reinforcement learning: Towards optimal sample complexity," *International Conference on Machine Learning (ICML)*, 2022.

C3. Pratik Patil, **Yuting Wei**, Alessandro Rinaldo, Ryan Tibshirani, "Uniform consistency of cross-validation estimators for high-dimensional ridge regression," *International Conference on Artificial Intelligence and Statistics (AISTATS)*, **oral presentation** (10.5% of accepted papers), 2021.

C4. Jingyan Wang, Ivan Stelmakh, **Yuting Wei**, Nihar B. Shah, "Debiasing Evaluations That are Biased by Evaluations," *AAAI Conference on Artificial Intelligence*, 2021.

C5. Shicong Cen, **Yuting Wei**, Yuejie Chi, "Fast Policy Extragradient Methods for Competitive Games with Entropy Regularization," *Neural Information Processing Systems (NeurIPS)*, 2021.

C6. Gen Li, Yuxin Chen, Yuejie Chi, Yuantao Gu, **Yuting Wei**, "Sample-Efficient Reinforcement Learning is Feasible for Linearly Realizable MDPs with Limited Revisiting," *Neural Information Processing Systems (NeurIPS)*, 2021.

C7. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Softmax Policy Gradient Methods

Can Take Exponential Time to Converge," *Conference on Learning Theory (COLT)*, 2021.

C8. Gen Li, Changxiao Cai, Yuxin Chen, Yuantao Gu, **Yuting Wei**, Yuejie Chi, "Tightening the Dependence on Horizon in the Sample Complexity of Q-Learning ," *International Conference on Machine Learning (ICML)*, 2021.

C9. Yue Li, Ilmum Kim, **Yuting Wei**, "Randomized Tests for High-Dimensional Regression: A More Efficient and Powerful Solution," *Neural Information Processing Systems (NeurIPS)*, 2020.

C10. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Breaking the Sample Size Barrier in Model-Based Reinforcement Learning with a Generative Model," *Neural Information Processing Systems (NeurIPS)*, 2020.

C11. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Sample Complexity of Asynchronous Q-Learning: Sharper Analysis and Variance Reduction," *Neural Information Processing Systems (NeurIPS)*, 2020.

C12. Chen Dan, **Yuting Wei** and Pradeep Ravikumar, "Sharp Statistical Guarantees for Adversarially RobustGaussian Classification," *International Conference on Machine Learning (ICML)*, 2020.

C13. **Yuting Wei**, Fanny Yang and Martin Wainwright, "Early Stopping for Kernel Boosting Algorithms: A General Analysis with Localized Complexities," *Neural Information Processing Systems (NeurIPS)*, **spotlight presentation**, 2017.

C14. **Yuting Wei** and Martin Wainwright, "Sharp Minimax Bounds for Testing Discrete Monotone Distributions," *International Symposium on Information Theory (ISIT)*, 2016.

Short Courses & Tutorials

S1. "Statistical and Algorithmic Foundations of Reinforcement Learning," CE Course, *Joint Statistical Meetings (JSM) 2023*, cotaught with Yuejie Chi and Yuxin Chen, Aug. 2023.

S2. "Reinforcement Learning: Fundamentals, Algorithms, and Theory," *International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2022*, cotaught with Yuejie Chi and Yuxin Chen, May. 2022.

S3. "Statistical and Algorithmic Foundations of Reinforcement Learning," *ICSA 2021 Applied Statistics Symposium Short Course*, cotaught with Yuejie Chi, Yuxin Chen and Zhengyuan Zhou, Sep. 2021.

Invited Talks

T1. "A non-asymptotic framework for the approximate message passing algorithm," *Statistics Colloquium, University of Chicago*, Mar. 2023.

T2. "A non-asymptotic framework for the approximate message passing algorithm," *Applied Probability and Risk (APR) Seminar, Columbia University*, Mar. 2023.

T3. "A non-asymptotic framework for approximate message passing in spiked models," *Statistics Annual Winter Workshop on Modern Computational Statistics, University of Florida, Gainesville*, Jan. 2023.

T4. "A non-asymptotic framework for approximate message passing in spiked models," *Computational and Methodological Statistics, London*, Dec. 2022.

T5. "Modern perspectives in high-dimensional statistics: Two recent stories," *Statistics Seminar, Tianyuan Mathematical Center in Central China, Wuhan*, Dec. 2022.

T6. "Modern perspectives in high-dimensional statistics: Two recent stories," *Statistical challenges and new solutions for large-scale complex data, Kunming, virtual*, Nov. 2022.

T7. "On the effectiveness and ineffectiveness of policy optimization for reinforcement learning," *Seminar of Statistics and Data Science, Hong Kong University of Science and Technology*, Oct. 2022.

T8. "Settling the sample complexity of model-based offline reinforcement learning," *Informs Annual Meeting, Indianapolis*, Oct. 2022.

T9. "Settling the sample complexity of model-based offline reinforcement learning," *Asilomar Conference on Signals, Systems, and Computers*, Oct. 2022.

T10. "A non-asymptotic framework for approximate message passing in spiked models," *SIAM Conference on Mathematics of Data Science (MDS22), San Diego*, Sep. 2022.

T11. "A non-asymptotic framework for approximate message passing in spiked models," *Probability Seminar, UC Davis*, Sep. 2022.

T12. "Beyond $o(\log n/\log\log n)$ iterations: A non-asymptotic framework for approximate message passing in spiked models," *Joint Statistical Meetings*, Washington, D.C., Aug. 2022.

T13. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Bernoulli Young Researchers Workshop*, virtual, Jul. 2022.

T14. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *ICSA China Conference*, virtual, Jul. 2022.

T15. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Youth in High-Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference*, Jun. 2022.

T16. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *International Conference on Econometrics and Statistics (EcoSta)*, Jun. 2022.

T17. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *New Advances in Statistics and Data Science, Hawaii*, May. 2022.

T18. "Beyond $o(\log n/\log\log n)$ Iterations: Non-asymptotic Analysis for Approximate Message Passing," *ICMS workshop, Edinburgh*, May. 2022.

T19. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Statistics Seminar, Cambridge University*, May. 2022.

T20. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Probabilitas Seminar, Harvard University*, Mar. 2022.

T21. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *52nd Annual Conference on Information Systems and Sciences (CISS) conference, Princeton*, Mar. 2022.

T22. "Exponential lower bounds and fast convergence for policy optimization," *Martin Wainwright's group, UC Berkeley*, Nov. 2021.

T23. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Workshop on Seeking Low Dimensionality in Deep Neural Networks (SLowDNN)*, Nov. 2021.

T24. "Breaking the sample size barrier in reinforcement learning," *Stochastics and Statistics Seminar, MIT*, Oct. 2021.

T25. "Bridge over troubled water: Lower bounds and fast convergence for policy optimization," *Informs Annual Meeting*, Oct. 2021.

T26. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent" *Peter Bartlett's group, UC Berkeley*, Oct. 2021.

T27. "Sample complexity of Q-learning: Sharper analysis and minimax optimality," *ICSA 2021 Symposium*, Sep. 2021.

T28. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Duke University*, Sep. 2021.

T29. "Breaking the sample size barrier in reinforcement learning," *Joint Statistical Meetings (JSM)*, Aug. 2021.

T30. "Softmax policy gradient methods can take exponential time to converge," *Conference on Learning Theory (COLT)*, Aug. 2021.

T31. "Statistical Inference over Convex Cones," *SIAM Conference on Optimization*, July. 2021.

T32. "Modern statistical perspectives in reinforcement learning and early stopping," *Statistics Seminar,*

Harvard University, Feb. 2021.

T33. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Wharton Statistics Seminar, University of Pennsylvania*, Feb. 2021.

T34. "Modern Statistical Perspectives in Reinforcement Learning and Early Stopping," *Statistics Seminar, Department of Data Sciences and Operations, University of Southern California*, Feb. 2021.

T35. "Breaking the Curse of Dimensionality: From Sparse Regression to Kernel Boosting," *Statistics Seminar, Stern School of Bussiness, New York University*, Jan. 2021.

T36. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Yale University*, Jan. 2021.

T37. "Breaking the Sample Size Barrier in Statistical Inference and Reinforcement Learning," *Wilks Statistics Seminar, ORFE, Princeton University*, Dec. 2020.

T38. "Breaking the Sample Size Barrier in Statistical Inference and Reinforcement Learning," *Statistics Seminar, Rutgers University*, Dec. 2020.

T39. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *Informs Annual Meeting*, Nov. 2020.

T40. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *Richard Samworth's group, Cambridge University*, Nov. 2020.

T41. "Reliable hypothesis testing paradigms in high dimensions," *Columbia University, Statistics Seminar*, Oct. 2020.

T42. "Reliable hypothesis testing paradigms in high dimensions," *Young Data Science Researcher Seminar Zurich, ETH*, Oct. 2020.

T43. "Reliable hypothesis testing paradigms in high dimensions" *Michigan State University, Statistics Seminar*, Oct. 2020.

T44. "The Lasso with general Gaussian design with application to hypothesis testing," *Joint Statistical Meeting*, Aug. 2020.

T45. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *TBSI Workshop on Learning Theory*, Shenzhen, Jul. 2020.

T46. "Understanding the distribution of the Lasso and its applications," *STATML Group, Carnegie Mellon University,* Mar. 2020.

T47. "A geometric perspective on hypothesis testing," *ICSA International Conference, Hangzhou, China,* Dec. 2019.

T48. "Towards a better understanding of the regularization in kernel learning," *TBSI Workshop on Learning Theory*, Shenzhen, Dec. 2019.

T49. "Towards a better understanding of the regularization in kernel learning," *Statistics Seminar, Cambridge University,* Nov. 2018.

T50. "Towards a better understanding of the regularization in kernel learning," *Big Data and Computational Social Science Lecture, University of British Columbia,* Mar. 2019.

T51. "Early stopping for gradient type algorithms," *Computational and Methodological Statistics, University of Pisa, Italy,* Dec. 2018.

T52. "The geometry of hypothesis testing over convex cones," *52nd Annual Conference on Information Systems and Sciences (CISS) conference, Princeton,* Mar. 2018.

T53. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, Carnegie Mellon University,* 2018.

T54. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, Stanford University,* 2018.

T55. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, University of Michigan,* 2018.

T56. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, University of Pennsylvania,* 2018.

T57. "Shape-constrained methods: Inference, applications, and practice," *Banff International Research Station for Mathematical Innovation and Discovery (BIRS), Canada,* 2018.

T58. "Early stopping for kernel boosting algorithms," *Neural Information Processing Systems (NIPS) conference, Long Beach,* 2017.

T59. Student talk for Oberwolfach Workshop "Statistical Recovery of Discrete, Geometric and Invariant Structures", 2017.

T60. "Sharp minimax bounds for testing discrete monotone distributions," *International Symposium on Information Theory (ISIT) conference, Barcelona,* 2016.

T61. "Adaptive estimation of planar convex sets," *Berkeley Statistics Annual Research Symposium (BSTARS),* 2016.

T62. "Sharp minimax bounds for testing discrete monotone distributions." *Stanford-Berkeley Joint Colloquium,* 2016.

Other Invited Workshops

- Statistics annual winter workshop on modern computational statistics. University of Florida, Gainesville, 2023.

- Semester program on computational complexity of statistical inference (long-term participant). Simons Institute, Berkeley, CA, 2021.

- Workshop on statistics meets machine learning. Oberwolfach, Germany, 2020.

- Workshop on statistical recovery of discrete, geometric and invariant structures. Oberwolfach, Germany, 2017.

Ph.D. Students & Postdocs

1. Wei Fan, Ph.D. in Statistics, 2022 -
2. Weichen Wu, Ph.D. in Statistics, 2019 - (co-advised with Alessandro Rinaldo)
3. Hong Hu, Postdoctoral researcher, 2022 - (co-advised with Yuxin Chen)
4. Gen Li, Postdoctoral researcher, 2021 - (co-advised with Yuxin Chen)
5. Yue Li, Ph.D. in Statistics, 2022 (co-advised with Kathryn Roeder)
6. Minshi Peng, Ph.D. in Statistics, 2021 (co-advised with Kathryn Roeder)

Doctoral Committees

1. Pratik Patil (CMU, Statistics & Machine Learning)
2. Chen Dan (CMU, Computer Science)
3. Yufei Yi (CMU, Statistics)

Teaching

University of Pennsylvania:

- STAT430: Probability, Spring 2022

Carnegie Mellon University:

- 36-747: Mathematics for High Dimensional Data: A statistical viewpoint, Spring 2021
- 36-748: Mathematics for High Dimensional Data: An optimization viewpoint, Spring 2021
- 36-225: Introduction to Probability, Fall 2020
- 36-741: Statistics meets Optimization: Iterative sketching methods, Fall 2019
- 36-742: Statistics meets Optimization: Approximate message passing algorithms, Fall 2019

Stanford University:

- Statistics 314: Advanced Statistical Theory, Spring 2019
- Statistics 206: Applied Multivariate Statistical Analysis, Winter 2018

University and Department Service

- Statistics departmental postdoc selection committee at UPenn statistics, 2021 - 2023
- Admission committee at UPenn statistics, 2021 - 2023
- Statistics department seminar co-organizer at CMU, 2019 - 2021
- Faculty senate member at CMU, 2019 - 2021
- Statistics department seminar co-organizer at Stanford, 2018 - 2019

Selected Professional Service

Grant reviewer: National Science Foundation

Reviewer for journals: Annals of Statistics, Journal of Machine Learning Research, Journal of the American Statistical Association, Biometrika, IEEE Transactions on Information Theory, Statistica Sinica, Electronic Journal of Statistics, Operations Research, SIAM Journal on Mathematics of Data Science, Journal of the Korean Statistical Society

Reviewer for conferences: International Conference on Artificial Intelligence and Statistics (AIS-TATS), IEEE International Symposium on Information Theory (ISIT), Conference on Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), Annual Conference on Learning Theory (COLT), Asian Conference on Machine Learning (ACML)

Conferences and Workshop Organization

- Co-organizer of the Oberwolfach Mini-workshop: Mathematical Foundations of Robust and Generalizable Learning, Germany, Oct 2022
- Session chair for 2022 INFORMS Annual Meeting, Indiana, Oct 2022.
- Session chair for 56th Conference for Information Sciences and Systems (CISS), Princeton, Mar 2022.
- Co-organizer for the AI panel of 2021 U8 World Innovation Summit, Sep 2021.